

Tilburg University

Perceptual effects of linguistic category priming

IJzerman, H.; Regenberg, N.; Saddlemeier, J.; Koole, S.L.

Published in:
Acta Psychologica

DOI:
[10.1016/j.actpsy.2015.01.008](https://doi.org/10.1016/j.actpsy.2015.01.008)

Publication date:
2015

Document Version
Publisher's PDF, also known as Version of record

[Link to publication in Tilburg University Research Portal](#)

Citation for published version (APA):
IJzerman, H., Regenberg, N., Saddlemeier, J., & Koole, S. L. (2015). Perceptual effects of linguistic category priming: The Stapel and Semin (2007) paradigm revisited in twelve experiments. *Acta Psychologica*, 157, 23-29. <https://doi.org/10.1016/j.actpsy.2015.01.008>

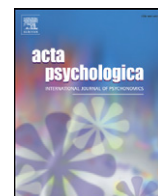
General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



Perceptual effects of linguistic category priming: The Stapel and Semin (2007) paradigm revisited in twelve experiments



Hans Ijzerman^{a,b,*}, Nina F.E. Regenberg^{a,1}, Justin Saddlemyer^{c,3}, Sander L. Koole^{a,4}

^a VU University, De Boelelaan 1105, Amsterdam 1081 HV, The Netherlands

^b Utrecht University, Campusplein 1, Utrecht 3584 ED, The Netherlands

^c VU University, De Boelelaan 1105, Amsterdam 1081 HV, The Netherlands

ARTICLE INFO

Article history:

Received 6 December 2013

Received in revised form 13 January 2015

Accepted 14 January 2015

Available online 19 February 2015

PsycINFO codes:

3040

2340

2720

Keywords:

Linguistic relativity

Priming

Linguistic Category Model

Referential

Replication

Data fraud

ABSTRACT

Linguistic category priming is a novel paradigm to examine automatic influences of language on cognition (Semin, 2008). An initial article reported that priming abstract linguistic categories (adjectives) led to more global perceptual processing, whereas priming concrete linguistic categories (verbs) led to more local perceptual processing (Stapel & Semin, 2007). However, this report was compromised by data fabrication by the first author, so that it remains unclear whether or not linguistic category priming influences perceptual processing. To fill this gap in the literature, the present article reports 12 studies among Dutch and US samples examining the perceptual effects of linguistic category priming. The results yielded no evidence of linguistic category priming effects. These findings are discussed in relation to other research showing cultural variations in linguistic category priming effects (Ijzerman, Saddlemyer, & Koole, 2014). The authors conclude by highlighting the importance of conducting and publishing replication research for achieving scientific progress.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Linguistic relativity refers to the idea that language shapes, or even determines, the way its speakers view the world (Sapir & Swadesh, 1946; Von Humboldt, 1843; Whorf, 1956). In its most radical form, linguistic relativity implies that language can be equated with thought. This extreme version of linguistic relativity has been clearly refuted by empirical observations (for discussions, see Gumpertz & Levinson, 1996; Pinker, 1994). Nevertheless, weaker forms of linguistic relativity may still apply, in which language influences (rather than fully determines) thought (Boroditsky, 2003; Gentner & Goldin-Meadow, 2003; Winawer et al., 2007). For instance, speakers of different languages are

able to perceive the same elementary colors of the spectrum, but having separate linguistic terms for perceptually close colors (e.g., different shades of blue) does allow speakers to differentiate such colors more easily (Winawer et al., 2007). Likewise, language may determine whether speakers use a relative frame of reference (to the left or right of, used by most English speakers) or an absolute frame of reference (north-west-south-east, Brown & Levinson, 1993; Majid, Bowerman, Kita, Haun, & Levinson, 2004) to orient themselves. Furthermore, language may shape whether people ground their sense of time in distance metaphors (e.g., “the time that lies ahead”) or in volume metaphors (e.g., “oceans of time”; e.g., Boroditsky, 2001; for an overview, see, Casasanto, 2008). In sum, although language cannot be equated with thought, a weaker form of linguistic relativity is still likely to be valid, such that language influences thought.

A recent development in linguistic relativity research has been the introduction of the linguistic category priming paradigm (Semin, 2008). Specific linguistic terms are recurrently paired with specific situations (Casasanto, 2008). As a result, linguistic terms may form associative links with cognitive processes. Because these associative links are stored in memory, they may be activated or “primed” whenever people encounter the relevant linguistic terms. One relevant dimension in this regard is linguistic abstraction. People can represent actions at different levels of linguistic abstraction (Semin & Fiedler, 1988; see also Semin,

* Corresponding author at: VU University, De Boelelaan 1105, Amsterdam 1081 HV, The Netherlands.

E-mail addresses: hijzerman@gmail.com (H. Ijzerman), nina@regenber.org (N.F.E. Regenberg), justin.saddlemyer@kuleuven.be (J. Saddlemyer), s.l.koole@vu.nl (S.L. Koole).

¹ Ijzerman and Regenberg contributed equally to this manuscript and share first authorship; authorship order for these two authors was determined in mutual agreement.

² Department of Social Psychology, School of Social and Behavioral Sciences, Tilburg University, The Netherlands.

³ Research Centre for Marketing and Consumer Science, Katholieke Universiteit, Leuven.

⁴ Department of Clinical Psychology, VU University Amsterdam, The Netherlands.

2001). A verb represents an action at a more concrete level (“Harry hits Peter”) than the same action that is described by an adjective (“Harry is aggressive”).⁵ Because abstract information is more general, it may be associated with global perceptions, whereas concrete information may become with local perceptions (Liberman, Trope, & Stephan, 2007). Consequently, priming verb may elicit a focus on local details (i.e., the trees), while priming adjectives may elicit a focus on the global whole (i.e., the forest).

Stapel and Semin (2007) reported four experiments using the linguistic category priming paradigm that confirmed the predicted effects of the different linguistic categories (verbs versus adjectives) on perceptual focus. Unfortunately, it later turned out that Stapel had fabricated the data (see Levelt Committee, Noort Committee & Drenth Committee, 2012; note that Semin was unaware of Stapel’s deception). Although this turn of events was deeply disturbing, fabricated data cannot speak to the true scientific validity (or invalidity) of a paradigm. Thus, priming linguistic categories may still have genuine and theoretically meaningful effects. Because the paradigm may yet contribute to the linguistic relativity debate, it remains important to investigate the effects of linguistic category priming.

Before Stapel was exposed, three of the present authors (IJzerman, Saddlemyer, and Koole, IJSK) were preparing a report of studies on the effects of priming different linguistic categories among Brazilian participants in 2009 (effects that are still subject to further investigation), for which they had planned to use Stapel and Semin’s (2007) data for cross-cultural comparison.⁶ Due to the invalidation of the Stapel and Semin report, the original plan of IJSK was no longer feasible. However, IJSK learned that Regenberg had conducted nine close and conceptual “replications” in the Netherlands between 2007 and 2009 (see Brandt et al., 2014, on the distinction between close and conceptual replication). As noted, the part of this paper based on Nina Regenberg’s research was conducted under the supervision of Gün R. Semin. After approximately 2 years of fruitless attempts to replicate and extend what was reported by Stapel and Semin (2007) this project was aborted. Gün R. Semin has no objections that this work is being used in the current publication (personal communication). In addition, IJzerman ran three experiments himself. Together, Regenberg and IJzerman conducted twelve studies using the linguistic category paradigm. Because these experiments represent the first genuine empirical research on this paradigm, they seem worthy of publication in their own right. We therefore report these twelve experiments here. The experiments provide a systematic series of empirical tests of the hypotheses Stapel and Semin (2007) had originally purported to test. Instead of discussing each experiments in detail, the present article provides a global outline of the main parameters that we varied in implementing linguistic category priming paradigm. In addition, we provide a meta-analytic summary of the results in this report. We refer readers who are interested in fuller study descriptions and results to our Open Science Framework webpage (<https://osf.io/f3kdu/>).⁷

2. Methods

2.1. Overview

The general idea behind the LCP paradigm is that participants are primed with different linguistic categories, which is followed by assessments of different manifestations of perceptual focus. These assessments were based on the assumption that priming abstract linguistic categories (adjectives) should lead to a more abstract perceptual focus, whereas concrete linguistic categories (action verbs) should lead to a more concrete perceptual focus. In the following paragraphs, we lay out the various ways in which we primed linguistic categories and how we measured perceptual focus.

2.2. Priming

Our most frequent method of priming linguistic categories was the scrambled sentence task (our Experiments 1–3, 6–9, 10–12). In this procedure, participants are typically given a jumble of words with which they were asked to construct a meaningful sentence (see our OSF page for a relevant example). By varying the contents of these word assortments, it is possible to prime participants with different linguistic categories. Thus, in the *adjective* condition, the word groups mostly contained adjectives (e.g., “aggressive,” “friendly,” “humble”). In the *verb* condition, these word groups mostly contained action verbs (e.g., “punch,” “help,” “swim”). The verb and adjective conditions both also contained neutral word groups as fillers.

A second priming method that we used was to ask participants to read priming sentences on a screen (our Experiments 4 and 5). Finally, a third method we employed was a subliminal priming task (our Experiment 8). In the latter procedure, participants were shown words from different linguistic categories (depending on condition) in alternating corners of the screen for a period of 60 ms, followed by a 60 ms mask.

2.3. Measurements

To measure perceptual processes, a first method we used was the framed line task (our Experiments 1, 2, and 3). In this task, participants were shown a box with a line of varying length suspended from the top edge (located half-way through the square). In the original variation of the task, participants are shown a second square of different dimensions (without the line) and are asked to recreate either the length of the first line (independent of the square), or recreate the line in proportion to the square (see, e.g., Fig. 1; Kitayama et al., 2003). In our experiments, we used only the absolute task. Thus, we could determine whether participants primed with different linguistic categories were prone to integrate irrelevant spatial information (the square) into their line judgments. Based on Stapel and Semin’s (2007) predictions, we predicted that priming adjectives would promote holistic perception and lead people to make more errors by incorporating the irrelevant information from the square. Conversely, priming verbs should activate a more concrete, detailed mindset, leading to better performance on the absolute version of the framed line task.⁸

⁵ Beyond the descriptive action verb, Semin and associates have also distinguished between interpretative/State Action Verbs and State Verbs. We focus on the distinction between DAVs and ADJs here because they constitute the two extremes and are hence most relevant for the discussion on concrete and abstract language. Moreover, research on linguistic category priming has so far mainly been concerned with the comparison between verbs and adjectives.

⁶ Despite repeated requests from the present authors during 2008–2010, Stapel did not provide the data from Stapel and Semin (2007).

⁷ Due to this data being collected several years ago, and our general expectation that we would be unable to report these studies, not all of the study materials have been retained. We have included the materials of the only study that was conducted in English on our webpage at the Open Science Framework.

⁸ In our Experiment, 3 we also measured similarity, attractiveness, height, age, and weight of in- and out-groups. We report the results of any auxiliary analyses on our webpage at the Open Science Framework.

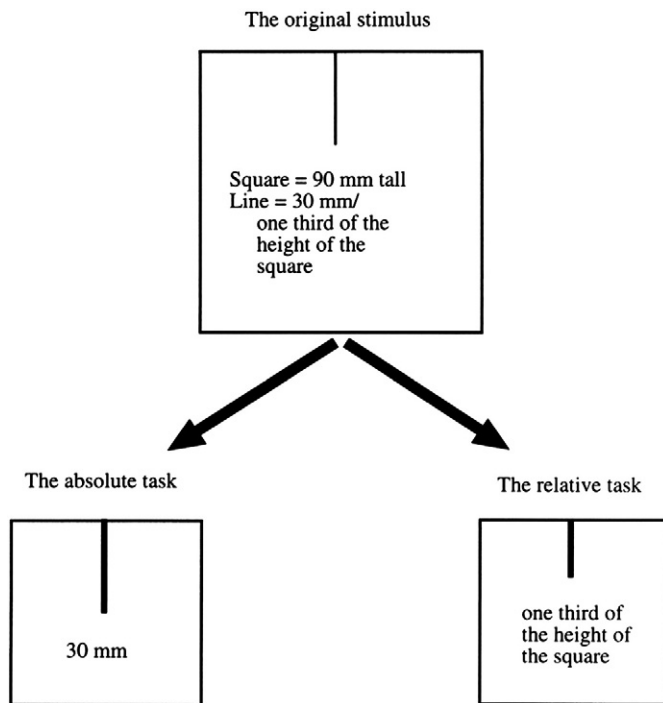


Fig. 1. Example of the framed line task. Participants were shown a square frame with a vertical line and asked to draw a line in a new square of the same or different size. The line was to be drawn so that it was identical to the first line in absolute length (absolute task) or so that the proportion between the length of the line and the height of its frame was identical to the proportion between the line and frame in the original stimulus (relative task; taken from Kitayama, Duffy, Kawamura, & Larsen, 2003).

Another task we used was a perceptual focus task, modelled after Kimchi and Palmer (1982; see Ijzerman & Semin, 2009; Schilder, Ijzerman, & Denissen, 2014; our Experiments 4, 5, and 11). This task consisted of a target object, which participants had to compare to two alternative figures and judge which figure they found to be most similar to the target (see e.g., Fig. 2). Like with the framed line task, one could

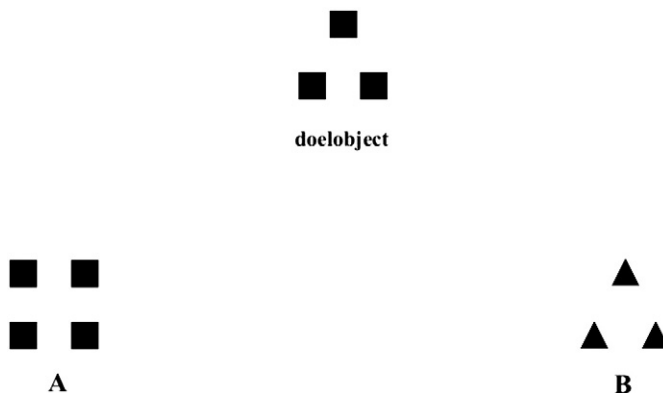


Fig. 2. Example of an item used in the perceptual focus task modelled after Kimchi and Palmer. The target object (in Dutch “doelobject”) is a triangle (pattern) consisting of three squares. From the line of reasoning presented in this paper, Figure A represents the global perspective (same pattern), while Figure B represents the local perspective (same individual properties).

expect those primed with verbs to have a more local perceptual focus, and thus to focus on independent properties of the images. Two other tasks we used to test our hypothesis are the Stroop task (our Experiments 6, 7, and 8), the Flanker task (Experiment 9), and the categorization task (our Experiments 10 and 12). In the former, participants were presented with color names with either incongruent or congruent font colors. By examining how participants perform on incongruent trials, one can infer “Stroop interference”; that is, how irrelevant information contained in the text influences responses. Based on Stapel and Semin’s (2007) ideas, we expected that priming participants with verbs (vs. adjectives) would lead to greater Stroop interference (a comparable line of reasoning would go for the Flanker task). In the latter experiment, we examine categorization by presenting participants with an item (i.e., a camel) and a category (i.e., a vehicle) and asking participants to indicate on a seven point scale how much they think the item belongs in the category. Using the examples provided in the brackets, one would expect that, priming with adjectives, at an abstract level, a camel could be construed as a vehicle, whereas this may less frequently be judged as such if the participant was primed with a verb (and would have a concrete focus).

3. Results

For ease of interpretation, we conducted a random-effects meta-analysis using the metaphor package in R (Viechtbauer, 2010) on our results in order to determine any effect of linguistic category priming across our between-subjects linguistic priming studies (see Table A1 for details of the methods). The average estimated standardized mean difference was not significant (Cohen’s d CI = 95–0.1603 0.2185, p = .76). The test for heterogeneity again was not significant (Q = 10.42, p = .24), which indicates that studies were comparable across dependent variables. Our 9 between-subjects studies examining the same effects of linguistic category priming amongst Dutch and American participants thus failed to demonstrate any traces of an effect (for our forest plot, see Fig. 3).

4. Discussion

I confabulated research and reported it. Colleagues and journal editors reviewed it critically. They published my work, and later it turned out that researchers in another city and in another country had done similar research. They had found the same results. My confabulated study was replicated. What appeared logical to me and what purely existed in my imagination turned out to be true.

In his autobiography *Derailment* (2012, p. 177), Stapel attributed some uncanny powers to himself. According to Stapel, his fabricated “studies” were replicated in actual experiments by other researchers. Thus, merely concocting a plausible-sounding finding was sufficient for the finding to become confirmed by experimental research. Stapel’s magical view of psychological science is strongly refuted by our twelve studies on linguistic category priming. Most of these studies were conducted before Stapel was exposed, when we ourselves were still convinced of the validity of the linguistic category priming effects reported by Stapel and Semin (2007). Nevertheless, the data proved us wrong: Our twelve studies, considered singly or in combination, yielded no evidence that linguistic category priming influences perceptual processing. (Indeed, the one significant effect that we observed was in the opposite direction of the proposed hypothesis.) The present research is thus a potent reminder that the outcomes of experimental psychology are grounded in empirical reality, no matter what researchers might expect or hope to find.

Many of the replication studies that are described in this article were conducted between 2007 and 2009 by Regenber, who consistently found null effects. At the time, these disconfirmatory findings were seen as “failed studies” that were not worthy of publication. In hindsight, it seems painfully clear that discarding null effects in this

Meta-analysis Verb/Adjective Priming

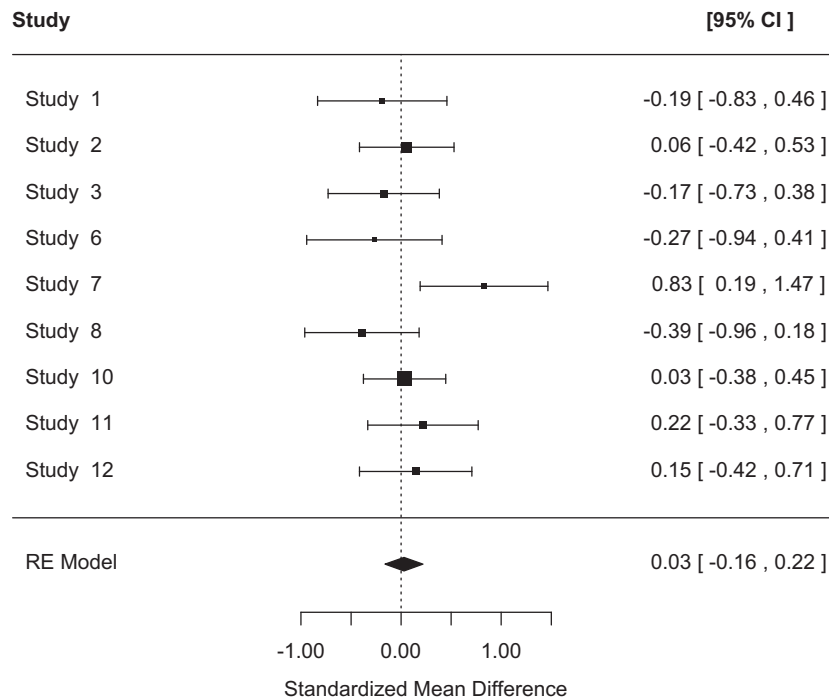


Fig. 3. Forest-plot of the between subjects studies in our paper. The model's effect size is a Cohen's d (with a CI 95 displayed for the model and for each study).

manner has hindered scientific progress. The present research thus constitutes a further argument for systematically conducting and reporting replication studies (see Brandt et al., 2014; Ijzerman, Brandt, & van Wolferen, 2013; Klein et al., 2014; Koole & Lakens, 2012). Publishing replication research, pre-registration, and archiving the data all contribute toward reducing redundancy in research efforts across different labs and to uncovering important contextual variations in empirical phenomena (for examples, see Schilder et al., 2014; Ong, Ijzerman, & Leung, 2014).

The present findings, combined with Stapel's data fraud, might be seen as the deathblow to the linguistic category priming paradigm, given that our data suggest that verb or adjective priming does not influence perceptual focus. However, in our view, discarding the whole idea of linguistic category priming would be throwing out the baby with the bathwater. One potential moderator for the usage of language is culture: We have recently proposed an alternative model of language use, which holds that verbs—in specific cultures—are more associated with relational processing than adjectives (Ijzerman, Saddlemyer, & Koole, 2014). According to this model, priming linguistic categories may well have consequences on people's perceptions, albeit in a different manner than was previously suggested by linguistic researchers (Semin, 2008). Our relational model of language use has so far been supported by a number of studies that we have collected amongst Brazilian participants. We have plans to further investigate these effects. Together with the present work, this alternative model of language use further suggests that perceptual effects of linguistic primes may differ meaningfully between countries and contexts.

If our foregoing considerations are valid, then the present studies highlight the need for a more sophisticated understanding of contextual influences on social-cognitive priming effects. From a situated

cognition perspective, several scholars have suggested that most social-cognitive priming effects do not exist as main effects but rather interact with people's goals, embodied states, and the specific situations people find themselves in (e.g., Smith & Semin, 2004; Ijzerman & Koole, 2011). Consequently, it remains to be seen whether the effect we have investigated does not exist, or whether it depends on identifying the right contexts and measurements for the linguistic category priming effects among Western samples (Cesario & Jonas, 2014).

Linguistic category priming may thus have significant influences on perception, even though these influences are likely to be highly complex and shaped by cultural (and/or linguistic) and situational constraints. This complexity and context sensitivity is consistent with a weak version of linguistic relativity, which posits that language is conditioned by (perceptual) experiences and hence *should* be reliant on cultural and situational constraints. Studying language and psychological processes in this way will initially require extra investments of time and resources, but the scientific benefits are sure to accumulate in later years.

Acknowledgments

This research was partially supported by a Veni grant of the Netherlands Organization for Scientific Research (NWO) (016.145.049) awarded to Hans Ijzerman and a consolidator grant from the European Research Council awarded to Sander L. Koole.

We are much indebted to Sylvie Kourilova, Mark Brandt, Gabriela Jiga-Boy, the editor Michael Kaschak, Debi Roberson, and one anonymous review for their insightful comments on earlier versions of this paper.

Appendix A

Table A1

Here we report all the methods and inferential statistics. Standardized means and standard deviations are presented.

Study	Sample	Procedure	DVs	Results
Experiment 1 (SS Study 1)	$N = 37$, 56.8% female	Participants completed the scrambled sentence task on the computer. They then completed the framed line task (absolute version) on the computer and a ten statements task on paper. The order of these two tasks was counterbalanced.	Framed line task (absolute). DV: error score between drawn and correct line length Expectation: Verbs priming leads to lower scores on the absolute task Meta-analysis approach: Scores were standardized and reversed (thus, verbs expected to have higher scores)	An independent-samples <i>t</i> -test revealed no significant difference between adjective and verb priming conditions; $t(35) = .57$, $p = .574$ on absolute error scores.
Experiment 2 (SS Study 1)	$N = 69$, 65.2% female	Participants completed the scrambled sentence task on paper. They then completed the framed line task (absolute and relative version, order counterbalanced) on the computer and a ten statements task on paper. The order of these two tasks was counterbalanced.	Framed line task (absolute). DV: error score between drawn and correct line length for both tasks Expectation: Verbs priming leads to lower scores on the absolute task Meta-analysis approach: Scores were standardized and reversed (thus, verbs expected with higher scores). We only included absolute scores in MA. Expectation: Adjective priming leads to lower scores on the relative task	An independent-samples <i>t</i> -test revealed no significant difference between adjective and action verb priming conditions; $t(67) = -.24$, $p = .81$ on absolute error scores. (An independent-samples <i>t</i> -test revealed a marginally significant difference between adjective and action verb priming conditions; $t(67) = 1.91$, $p = .06$ on relative error).
Experiment 3 (SS Studies 2 and 3)	$N = 50$, 64% female	Participants completed the scrambled sentence task on paper. They then made judgments of how close two displayed individuals seemed to each other and rated them individually on the features attractiveness, height, weight, and age. After a surprise recall of the scrambled sentence task items, participants completed the framed line task (absolute version) on the computer.	Framed line task (absolute). DV: error score between drawn and correct line length Expectation: Verbs priming leads to lower scores on the absolute task Meta-analysis approach: Scores were standardized and reversed (thus, verbs expected to have higher scores)	An independent-samples <i>t</i> -test revealed no significant difference between adjective and action verb priming conditions; $t(48) = .62$, $p = .54$ on absolute error scores.
Experiment 4 (SS Study 1)	$N = 24$, 68.2% female	Participants received an in-group priming. Participants, all students of VU University, wrote down five stereotypic traits of VU and of UvA (competing university) students. They then went on to a modified perceptual focus task. Before each trial they read a sentence on the screen. These were designed to either contain a verb or an adjective (language abstraction) and they described a positive or a negative behavior/trait (valence), which was congruent with the respective stereotype. There were four blocks of 24 trials. The perceptual focus task trials were identical in each block, but there were 120 unique sentences. Pop-up multiple choice questions about the sentences were included to insure that participants properly read them and between the blocks participants solved two math problems.	Perceptual focus task DV: choice for global vs. local option (coding: local = 1, global = 0, thus the DV represents the amount of locality) Expectation: Verbs priming leads to higher scores on the perceptual focus task Meta-analysis approach: Given design was substantially different (repeated measures, and interactions), we omitted this study from the meta-analysis.	A 2 (language abstraction) \times 2 (in-group/out-group) \times 2 (valence) repeated measures revealed no significant main effect of language on perceptual focus, ($F(1,21) = .00$, $p = 1.00$).
Experiment 5 (SS Study 1)	$N = 27$, 63 % female	Same as in Experiment 4 except that a response window was introduced for reading the sentences.	Perceptual focus task DV: choice for global vs. local option (same as in Experiment 4) Expectation: Verbs priming leads to higher scores on the perceptual focus task Meta-analysis approach: Given that the design was substantially different (within design), we omitted this study from the meta-analysis	A 2 (language abstraction) \times 2 (in-group/out-group) \times 2 (valence) repeated measures revealed no significant main effect of language on perceptual focus, ($F(1,26) = .24$, $p = .63$)
Experiment 6 (Conceptual Replication)	$N = 61$ (7 excluded for color blindness), 70.4% female. For this paper, we conducted analyses conducted on participants in adjectives and action verb conditions ($N = 34$)	Participants completed an SST, priming language abstraction in three conditions: ADJ, SV, and DAV. They then proceeded to a Stroop task consisting of five blocks of 30 trials each (10 congruent, 10 neutral, 10 incongruent): A fixation cross	Stroop task DV: Stroop interference on reaction times of correct responses (difference between incongruent and neutral trials, analysis on RT means within 2 SD of respective condition mean of correct responses)	An independent-samples <i>t</i> -test revealed no significant difference between adjective and action verb priming conditions; $t(32) = -.796$, $p = .43$ on reaction times.

(continued on next page)

Table A1 (continued)

Study	Sample	Procedure	DVs	Results
		appeared in the middle of the screen for 1 second, followed by either the word “blue,” the word “red,” or a row of x’s (XXXXX), which were either written in red or blue. Participants’ task was to indicate the word’s color (not its meaning) by pressing a button with their left or right hand, respectively.	Expectation: Verbs priming leads to greater Stroop interference Meta-analysis approach: Scores were standardized	
Experiment 7 (Conceptual Replication)	<i>N</i> = 48 (7 excluded for color blindness), 73.2% female	Participants completed a scrambled sentence task priming language abstraction in two conditions: adjectives and action verbs. They then proceeded to a paper-and-pencil ten statements task (not reported). Afterward, participants received a surprise recall task about the scrambled sentence task and then performed the Stroop task as in Experiment 6.	Stroop task DV: Stroop interference on reaction times of correct responses (analysis on means within 2 SD of respective condition mean of correct responses) Expectation: Verbs priming leads to greater Stroop interference Meta-analysis approach: Scores were standardized	An independent-samples <i>t</i> -test revealed a significant difference between adjective and action verb priming conditions; $t(39) = 2.70$, $p = .010$ on reaction times (counter to the expected direction).
Experiment 8 (Study 4)	<i>N</i> = 48, 62.5% female	Participants received a subliminal language priming: A fixation cross was continuously shown in the middle of the screen. In a varying interval of 2 to 7 seconds, words were shown in one of the four corners of the screen for 60 ms, immediately followed by a mask (“xqfbzrmqwgxb”) for another 60 ms. Participants then had to indicate whether the “flash” they saw appeared at the left or at the right side of the screen. Then the next trial started. For half of the participants, the prime words consisted of 40 adjectives and 20 neutral filler items (e.g., table, chair). The other half of the participants was presented with 40 action verbs and 20 neutral filler items. Subsequently, participants performed the Stroop task as in Experiment 6.	Stroop task DV: Stroop interference on reaction times of correct responses (analysis on means within 2 SD of respective condition mean of correct responses) Expectation: Verbs priming leads to greater Stroop interference Meta-analysis approach: Scores were standardized.	An independent-samples <i>t</i> -test revealed no significant difference between adjective and action verb priming conditions; $t(46) = -1.378$, $p = .18$ on reaction times.
Experiment 9 (Conceptual Replication)	<i>N</i> = 27, 81.5% female	Participants completed a scrambled sentence task priming language abstraction in two conditions: adjectives and action verbs. They then proceeded to a flanker task. On each trial, participants saw a row of five arrows or five lines. They had to respond to the middle arrow by pressing one of two buttons to indicate in which direction the arrow was pointing. The flanking arrows, two on each side, all pointed in the same direction which was either congruent or incongruent with the response stimulus. On a third type of trials the flanking stimuli were neutral lines without an implied direction. Furthermore, the row of stimuli appeared in one of three vertical locations: at the top, the center or the bottom of the screen. Participants did 12 neutral, 12 congruent, and 12 incongruent trials.	Flanker task DV: reaction times of correct responses (analysis on means within 2 SD of respective condition mean of correct responses) Expectation: Verbs priming leads to greater interference Meta-analysis approach: Given that the design was substantially different (within design), we omitted this study from the meta-analysis..	A 2 (language abstraction, between) \times 2 (trial type, within) \times 3 (position, within) repeated measures revealed a marginally significant main effect of language on flanker task performance, ($F(1,25) = 3.48$, $p = .074$). However, none of the interactions with trial type or position were significant ($p = .89$ and $.44$, respectively).
Experiment 10 (SS Study 2)	<i>N</i> = 91, 41.8% female	Participants completed an adjusted scrambled sentence task (SST) priming language abstraction in two conditions: adjectives and action verbs. They then proceeded to a category inclusion task. They had to indicate how representative they perceived a particular item to be of a category on a slider from definitely does not belong to this category to definitely belongs to this category.	Inclusiveness of categorization DV: Inclusiveness of categorization (greater inclusiveness equates to more global perceptual focus) Expectation: Verbs priming leads to lower inclusiveness scores Meta-analysis approach: Scores were standardized and reversed (thus, verbs expected to have higher scores)	An independent-samples <i>t</i> -test revealed no significant difference between adjective and verb priming conditions; $t(89) = -.168$, $p = .867$ on inclusiveness of categorization.

Table A1 (continued)

Study	Sample	Procedure	DVs	Results
Experiment 11 (Conceptual Replication)	N = 52, 75% female	Participants completed a scrambled sentence task priming language abstraction in two conditions: adjectives and action verbs. They then completed a 24-trial perceptual focus task	Perceptual focus task DV: choice for global vs. local option (coding was reversed, so we standardized the dependent variable, and calculated its negative score). Expectation: Verbs priming leads to higher scores on the perceptual focus task Meta-analysis approach: Scores were standardized	An independent-samples <i>t</i> -test revealed no significant difference between adjective and verb conditions; $t(49) = .79, p = .43$ on perceptual focus.
Experiment 12 (SS Study 2)	N = 49, 55.1% female	Participants completed a scrambled sentence task priming language abstraction in two conditions: adjectives and action verbs. They then proceeded to the category inclusion task.	Inclusiveness of categorization DV: Inclusiveness of categorization (greater inclusiveness equates to more global perceptual focus) Expectation: Verbs priming leads to lower inclusiveness scores Meta-analysis approach: Scores were standardized and reversed (thus, verbs expected to have higher scores)	An independent-samples <i>t</i> -test revealed no significant difference between adjective and verb conditions, $t(47) = -.52, p = .61$ on inclusiveness of categorization.

References

- Boroditsky, L. (2001). Does language shape thought? Mandarin and English speakers' conception of time. *Cognitive Psychology*, 43, 1–22.
- Boroditsky, L. (2003). *Linguistic relativity*. Encyclopedia of Cognitive Science.
- Brandt, M.J., IJzerman, H., Dijksterhuis, A., Farach, F.J., Geller, J., Giner-Sorolla, R., et al. (2014). The Replication Recipe: What makes for a convincing replication? *Journal of Experimental Social Psychology*, 50, 217–224.
- Brown, P., & Levinson, S.C. (1993). "Uphill" and "downhill" in Tzeltal. *Journal of Linguistic Anthropology*, 3, 46–74.
- Casasanto, D. (2008). Who's afraid of the Big Bad Whorf? Crosslinguistic differences in temporal language and thought. *Language Learning*, 58, 63–79.
- Cesario, J., & Jonas, K.J. (2014). Replicability and models of priming: What a resource computation framework can tell us about expectations of replicability. *Social Cognition*, 32, 124–136.
- Gentner, D., & Goldin-Meadow, S. (2003). Whither Whorf. In D. Gentner, & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and cognition* (pp. 3–14). Cambridge, MA: MIT Press.
- Gumpertz, J.J., & Levinson, S.C. (Eds.). (1996). *Rethinking Linguistic Relativity*. Cambridge University Press.
- Humboldt, W. (1843). Ueber das Entstehen der grammatischen Formen und deren Einfluss auf die Ideenentwicklung [On the development of grammatical forms and their influence on the development of ideas]. In *Gesammelte Werke* 3. (pp. 269–306) Berlin, Germany: Reimer.
- IJzerman, H., Brandt, M.J., & van Wolferen, J. (2013). Rejoice! In replication. *European Journal of Personality*, 27, 128–129.
- IJzerman, H., & Koole, S.L. (2011). From perceptual rags to metaphoric riches: Bodily, social, and cultural constraints on socio-cognitive metaphors. *Psychological Bulletin*, 137, 355–361.
- IJzerman, H., Saddlemyer, J., & Koole, S. L. (2014a). The role of verbs and adjectives in Brazilian samples. *Unpublished Dataset*.
- IJzerman, H., & Semin, G.R. (2009). The thermometer of social relations: Mapping social proximity on temperature. *Psychological Science*, 1214–1220.
- Kimchi, R., & Palmer, S.E. (1982). Form and texture in hierarchically constructed patterns. *Journal of Experimental Psychology: Human Perception and Performance*, 8, 521–535.
- Kitayama, S., Duffy, S., Kawamura, T., & Larsen, J.T. (2003). Perceiving an object and its context in different cultures: A cultural look at new look. *Psychological Science*, 14(3), 201–206.
- Klein, R.A., Ratliff, K.A., Vianello, M., Adams, R.B., Jr., Bahník, Š., Bernstein, M.J., et al. (2014). Investigating variation in replicability. *Social Psychology*, 45, 142–152.
- Koole, S.L., & Lakens, D. (2012). Rewarding replications: A sure and simple way to improve psychological science. *Perspectives on Psychological Science*, 7, 608–614.
- Levelt Committee, Noort Committee, & Drenth Committee (2012). Flawed science: The fraudulent research practices of social psychologist Diederik Stapel. (Retrieved June 28, 2013 from) https://www.commissielevelt.nl/wp-content/uploads_per_blog/commissielevelt/2013/01/finalreportLevelt1.pdf
- Liberman, N., Trope, Y., & Stephan, E. (2007). Psychological distance. *Social Psychology: Handbook of Basic Principles*, 2, 353–383.
- Majid, A., Bowerman, M., Kita, S., Haun, D., & Levinson, S.C. (2004). Can language restructure cognition? The case for space. *Trends in Cognitive Sciences*, 8, 108–114.
- Ong, L., IJzerman, H., & Leung, A.K.-y. (2014). Comfort food across cultures: A replication of Troisi & Gabriel. *Frontiers in Cognition*.
- Pinker, S. (1994). *The language instinct*. New York: William Morrow and Company.
- Sapir, E., & Swadesh, M. (1946). American Indian grammatical categories. *Word*, 2, 103–112.
- Schilder, J.D., IJzerman, H., & Denissen, J.J. (2014). *Physical warmth and perceptual focus: A replication of IJzerman and Semin (2009)*. (Available at SSRN 2378184).
- Semin, G.R. (2001). Language and social cognition. In A. Tesser, & N. Schwarz (Eds.), *Blackwell handbook of social psychology. Intra-Individual Processes*, Vol. 1. (pp. 159–180). Oxford: Blackwell.
- Semin, G.R. (2008). Language puzzles: A prospective retrospective on the Linguistic Category Model. *Journal of Language and Social Psychology*, 27, 197–209.
- Semin, G.R., & Fiedler, K. (1988). The cognitive functions of linguistic categories in describing persons: Social cognition and language. *Journal of Personality and Social Psychology*, 54, 558–568.
- Smith, E. R., & Semin, G. R. (2004). Socially situated cognition: Cognition in its social context. *Advances in experimental social psychology*, 36, 57–121.
- Stapel, D.A. (2012). *Ontsporing (Derailment)*. Amsterdam: Prometheus (Available online at <http://tilburgvirtueel.nl/images/stories/pdf/ontsporing.pdf>).
- Stapel, D.A., & Semin, G.R. (2007). The magic spell of language: Linguistic categories and their perceptual consequences. *Journal of Personality and Social Psychology*, 93, 23–33.
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36.
- Whorf, B.L. (1956). *Language, thought, and reality*. Cambridge, MA: MIT Press.
- Winawer, J., Witthoft, N., Frank, M.C., Wu, L., Wade, A.R., & Boroditsky, L. (2007). Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences*. 104. (pp. 7780–7785).